Annealed lower tails for the energy of a polymer.

Amine Asselah
Université Paris-Est
Laboratoire d'Analyse et de Mathématiques Appliquées
UMR CNRS 8050
amine.asselah@univ-paris12.fr

Abstract

We consider the energy of a randomly charged polymer. We assume that only charges on the same site interact pairwise. We study the lower tails of the energy, when averaged over both randomness, in dimension three or more. As a corollary, we obtain the *correct* temperature-scale for the Gibbs measure.

Keywords and phrases: random polymer, large deviations, random walk in random scenery, self-intersection local times.

AMS 2000 subject classification numbers: 60K35, 82C22, 60J25.

Running head: Lower tails for the energy a polymer.

1 Introduction

In this paper, we study the lower tails for the energy of a polymer. This complements a companion paper [1] dealing with the upper tails. Lower and upper tails are different stories, and the two papers are independent from each other, though they use the same model, and the same notations. Thus, our polymer is a linear chain of n monomers each carrying a random charge, and sitting sequentially on the positions of a symmetric random walk.

(i) The symmetric random walk on $\mathbb{Z}^d$ is denoted $\{S(n), n \in \mathbb{N}\}$. When $S(0) = z \in \mathbb{Z}^d$, its law is denoted $\mathbb{P}_z$.

(ii) The random field of charges is denoted $\{\eta(n), n \in \mathbb{N}\}$. The charges are centered i.i.d. with a finite forth moment. We denote by η a generic charge variable, and the charges' law is denoted by Q.

The monomers interact pairwise only when they occupy the same site on the lattice. The interaction produces an energy

$$H_n = \sum_{z\in\mathbb{Z}^d} \sum_{0\le i\ne j<n} \eta(i)\eta(j)\ \mathbb{I}\{S(i) = S(j) = z\}. \tag{1.1}$$

Our toy-model comes from physics, where it is used to model proteins or DNA *folding*. However, physicists' usual setting differs from ours by three main features. (i) Their polymer is usually *quenched*: a typical realization of the charges is fixed, and the average is over the walk. (ii) A short-range repulsion is included by considering random walks such as the self-avoiding walk or the directed walk. (iii) The averages are performed with respect to the the so-called Gibbs measure: a probability measure obtained from $\mathbb{P}_0$ by weighting it with $\exp(\beta H_n)$, with real parameter β. When β is positive, the Gibbs measure favors configuration with large energy; in other words, alike charges attract each other: this models *hydrophobic interactions*, where the effect of avoiding the water solvent is mimicked by an attraction among hydrophobic monomers. When β is negative, alike charges repel: this models Coulomb potential, and describes also the effective repulsion between identical bases of RNA. The issue is whether there is a *critical value* $\beta_c(n)$, such that as β crosses $\beta_c(n)$, a phase transition occurs. For instance, Garel and Orland [14] observed a phase transition as β crosses a $\beta_c(n) \sim 1/n$, from a collapsed shape to a random-walk like shape. Kantor and Kardar [15] discussed the quenched model for the case $\beta < 0$, that is when alike charges repel. Some heuristics (dimensional analysis on the continuum version) suggests that the *(upper) critical* dimension is 2: for $d \geq 3$, the polymer looks like a simple random walk, whereas when $d < 2$, its average end-to-end distance is n^ν with $\nu = \frac{2}{d+2}$. Let us also mention studies of Derrida, Griffiths and Higgs [11] and Derrida and Higgs [12]: both study the quenched Gibbs measure $\exp(-\beta H_n) d\tilde{\mathbb{P}}_0$, with $\beta > 0$, for a one dimensional directed random walk $\tilde{\mathbb{P}}_0$, and obtain evidence for a phase transition (a so-called weak freezing transition).

Our interest stems from recent mathematical works of Chen [8], and Chen and Khoshnevisan [10], dealing with central limit theorems for H_n. Chen [8] establishes also an annealed moderate deviation principle, under the additional assumption that $E[\exp(\lambda\eta^2)] < \infty$, for some $\lambda > 0$. More precisely, with the annealed law denoted P, $d \geq 3$, $n^{\frac{1}{2}} \ll \sqrt{n}\xi_n \ll n^{\frac{2}{3}}$, (for two positive sequences $\{a_n, b_n, n \in \mathbb{N}\}$, we say that $a_n \ll b_n$, when $\limsup \frac{\log(a_n)}{\log(b_n)} < 1$), X.Chen has obtained

$$\lim_{n\to\infty} \frac{1}{\xi_n^2} \log\left(P(\pm\frac{H_n}{\sqrt{n}} \geq \xi_n)\right) = -\frac{1}{2c_d}, \quad \text{where} \quad c_d = \sum_{n\geq 1} \mathbb{P}_0(S(n) = 0). \tag{1.2}$$

Our study complements the work [8]. We study the annealed probability that $\{-H_n > \xi_n\}$ for $\xi_n \geq n^{\frac{2}{3}}$. Also, we consider the simplest aperiodic walk: the walk jumps to a nearest neighbor site or stays still with equal probability.

As in [1], we rewrite the energy into a convenient form. For $z \in \mathbb{Z}^d$, and $n \in \mathbb{N}$, we call $l_n(z)$ the *local times*, and $\check{q}_n(z)$ the *local charges*. That is

$$l_n(z) = \sum_{k=0}^{n-1} 1\!\!1\{S(k) = z\}, \quad \text{and} \quad \check{q}_n(z) = \sum_{k=0}^{n-1} \eta(k)\, 1\!\!1\{S(k) = z\}.$$

We write $H_n = \sum_z \check{X}_n(z) + Y_n(z)$ with

$$\check{X}_n(z) = \check{q}_n^2(z) - l_n(z), \quad \text{and} \quad Y_n(z) = l_n(z) - \sum_{i=0}^{n-1} \eta(k)^2\, 1\!\!1\{S(k) = z\}.$$

Now,

$$Y_n = \sum_{z\in\mathbb{Z}^d} Y_n(z) = \sum_{i=0}^{n-1} \left(1-\eta^2(i)\right), \tag{1.3}$$

is a sum of centered independent random variables, and its large deviation asymptotic are well known (see below Remark 1.4). Thus, we focus on $\check{X}_n = \sum_{\mathbb{Z}^d} \check{X}_n(z)$.

Before presenting our lower tails estimates, we provide some heuristics.

Heuristics. Since we are interested in annealed estimates, note that

$$\check{X}_n \overset{\text{law}}{=} X_n := \sum_{z\in\mathbb{Z}^d} l_n(z)\left(\zeta_z(l_n(z))-1\right), \quad \text{where} \quad \zeta_z(n) = \left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_z(i)\right)^2, \tag{1.4}$$

where $\{\eta_z(i), z\in\mathbb{Z}^d,\ i\in\mathbb{N}\}$ is an i.i.d. sequence with $\eta_z(i)\sim\eta$, and we still denote its law with Q. Let us fix two lengths T_n and r_n, and an energy x_n, and estimate the cost of folding T_n-monomers in a ball of radius r_n, say $B(r_n)$, in order to realize

$$\sum_{z\in B(r_n)} l_n(z)\left(1-\zeta_z(l_n(z))\right) \ge x_n.$$

Note that necessarily $T_n \ge x_n$. Assume also that $T_n \gg |B(r_n)|$, so that we expect many monomers to pile up on each site of $B(r_n)$, and we further assume that the filling is uniform, that is

$$\forall z\in B(r_n), \quad l_n(z) \sim \frac{T_n}{|B(r_n)|}.$$

Then, the optimal scenario comes up as we equate the cost of the two constrains we are imposing. (i) We localize the walk a time T_n in a ball $B(r_n)$. This costs of the order of $\exp(-\kappa T_n |B(r_n)|^{-2/d})$. (ii) We require the charges to realize

$$\left\{\sum_{z\in B(r_n)} 1-\zeta_z(l_n(z)) \ge \frac{x_n|B(r_n)|}{T_n}\right\}. \tag{1.5}$$

Since, when we freeze the walk, the variables $\{1-\zeta_z(l_n(z)), z\in B(r_n)\}$ are independent, centered and with finite variance (if $E[\eta^4]<\infty$), the cost of (1.5) is

$$P\left(\sum_{z\in B(r_n)} 1-\zeta_z(l_n(z)) \ge \frac{x_n|B(r_n)|}{T_n}\right) \sim \exp\left(-\frac{x_n^2|B(r_n)|}{T_n^2}\right). \tag{1.6}$$

As we equate the two costs, we find

$$\frac{x_n^2|B(r_n)|}{T_n^2} = \frac{T_n}{|B(r_n)|^{2/d}} \Longrightarrow |B(r_n)|^{\frac{d+2}{d}} = \frac{T_n^3}{x_n^2}. \tag{1.7}$$

Thus, the heuristic discussion suggests that for some constant $c>0$

$$P(X_n \le -x_n) \sim \exp\left(-c x_n^{\frac{4}{d+2}} T_n^{\frac{d-4}{d+2}}\right). \tag{1.8}$$

Note that the exponent $\frac{d-4}{d+2}$ of T_n in (1.8) suggests that $d=3$ and $d>4$ have a distinct phenomenology. When $d=3$, the cheapest cost is reached when $T_n=n$: the polymer is *entirely folded* in a ball of volume $(\frac{n^3}{x_n^2})^{\frac{3}{5}}$. Also, the sum of local charges, $\check{q}_n$, over this domain performs a *moderate deviations*.

When $d>4$, the cheapest cost requires the smallest T_n, which is $x_n \le n$. Thus, the polymer is *partially folded*, and (1.8) implies that the volume of the ball is $x_n^{\frac{d}{d+2}}$. Also, on each site the *local charge* performs a *typical fluctuation*.

Our heuristics set the stage for the following mathematical statements.

Theorem 1.1 *Assume $d=3$, and $E[\eta^4]<\infty$. There are constants $a_0, c_3^{\pm}$ such that for $a_0 \le \xi_n < n^{1/3}$,*

$$\exp\left(-c_3^- \xi_n^{\frac{4}{5}} n^{\frac{1}{3}}\right) \le P(\check{X}_n \le -\xi_n n^{2/3}) \le \exp\left(-c_3^+ \xi_n^{\frac{4}{5}} n^{\frac{1}{3}}\right). \tag{1.9}$$

Moreover, we have the following description of the dominant strategy. For a constant a large enough,

$$\lim_{n\to\infty} P\left(|\{z\in\mathbb{Z}^d : \frac{\xi_n^{\frac{6}{5}}}{a} \le l_n(z) \le a\xi_n^{\frac{6}{5}}\}| \ge \frac{n}{a^4 \xi_n^{6/5}} \;\Big\|\; \check{X}_n \le -\xi_n n^{\frac{2}{3}}\right) = 1. \tag{1.10}$$

In dimension 4 and more, there are two regimes. In the following regime, the energy has the same behavior as in the moderate deviation regime, where the polymer is *unfolded*.

Theorem 1.2 *Assume $d\ge 4$, and $E[\eta^4]<\infty$. For any ϵ positive, choose any sequence $\{\xi_n\}$ with*

$$\xi_n \in [n^{1/6}, n^{(d/2)/(d+4)-\epsilon}].$$

There are $c_1, c_2>0$, such that for n large enough

$$\exp\left(-c_1\xi_n^2\right) \le P\left(\check{X}_n \le -\xi_n\sqrt{n}\right) \le \exp\left(-c_2\xi_n^2\right). \tag{1.11}$$

Moreover, for a constant A large enough

$$\lim_{n\to\infty} P\left(\sum_{z:\ l_n(z)\ge A} \check{X}_n(z) \le -\xi_n\sqrt{n}\right) = 0. \tag{1.12}$$

The second regime corresponds to a *partially folded polymer* as alluded to in the heuristic discussion.

Theorem 1.3 *Assume $d\ge 4$, and $n^{\frac{d+2}{d+4}} < \xi_n \le \xi n$ with $\xi<1$. For a constant c_d^-, and for any $\epsilon>0$,*

$$\exp\left(-c_d^- \xi_n^{\frac{d}{d+2}}\right) \le P\left(\check{X}_n \le -\xi_n\right) \le \exp\left(-\xi_n^{\frac{d}{d+2}} n^{-\epsilon}\right). \tag{1.13}$$

Remark 1.4 *The lower tail behavior of H_n depends on a competition between $\check{X}_n$ and Y_n whose upper tail behavior is given in Remark 2.2. Let us mention that if $\alpha \geq \frac{2d}{d+2}$, then the lower tails of H_n are identical to that of $\check{X}_n$. When $d \geq 4$, and $\alpha < \frac{2d}{d+2}$, then Y_n dictates the behavior of H_n: the correct speed for the lower tails of H_n is $\min(\xi_n^2/n,\ \xi_n^{\alpha/2})$. In $d = 3$, the correct speed for the lower tails of H_n is $\min(\xi_n^{4/5} n^{-1/5},\ \xi_n^{\alpha/2})$. Thus, as soon as $\alpha \geq 2$, the lower tails of H_n are identical to that of $\check{X}_n$.*

Remark 1.5 *The weakness in the upper bound in (1.13) (the artifact $n^{-\epsilon}$ in the exponent) reflects a deep technical gap in estimating the distribution of the size of level sets of the local times of the random walk. We state it as a conjecture.*

Conjecture 1.6 *Assume $d \geq 3$, and let $\{y_n, n \in \mathbb{N}\}$ be a sequence going to infinity, with $y_n^{1+d/2} \leq n$. Then, there is $\kappa_d > 0$ (independent on n) such that*

$$\mathbb{P}_0 \left(| \{ z : \ l_n(z) \geq y_n \} | \geq y_n^{d/2} \right) \leq \exp(-\kappa_d y_n^{d/2}). \tag{1.14}$$

One way to understand the difficulty of (1.14) is to see that the number of possible regions of volume $y_n^{d/2}$ inside $[-n, n]^d$ exceeds $\exp(\kappa y_n^{d/2})$, for any $\kappa > 0$.

We give now an elementary application of Theorem 1.1 to the study of annealed Gibbs measure in dimension three. For simplicity, we further assume that $\eta \in \{-1, 1\}$, so that $H_n = \check{X}_n$. The annealed Gibbs measure is the following probability measure: for $\beta > 0$, we set

$$dP^-_{n,\beta} = \frac{\exp(-\beta H_n) dP}{Z_n^-(\beta)} \quad \text{where} \quad Z_n^-(\beta) = E\left[\exp(-\beta H_n)\right]. \tag{1.15}$$

The normalizing constant $Z_n^-(\beta)$ is called partition function. The measure $P^-_{n,\beta}$ favors configurations with large values of $-H_n$, so that it forces local charges to neutralize. When dealing with the Gibbs measure, the issue is to find the *correct* temperature-scaling for which a phase-transition occurs. Indeed, the interesting biological phenomenon which motivates polymer modelling is *folding*, that is the process of going from a (transient) random-walk shape to a globular-looking shape, under the tuning of temperature, or salt-concentration. Thus, we expect a critical parameter $\beta_c(n)$ (which might scale with the polymer size), such that for $\beta > \beta_c(n)$, typical polymers are globular-like looking, whereas when $\beta < \beta_c(n)$, typical polymers look like typical random walk trajectories.

Biskup and König [6] (see also Buffet and Pulé [7]) obtain results and some heuristics on the *annealed* Gibbs measure (i.e. averaged over both randomness). They use that when freezing the random walk, and averaging over charges

$$E_Q[e^{-\beta H_n}] = c_n \exp(-\sum_{z \in \mathbb{Z}^d} V(l_n(z))) \quad \text{where for } x \text{ large} \quad V(x) \sim \frac{1}{2} \log(1 + 2\beta x), \tag{1.16}$$

where $\beta > 0$ and c_n is a constant. When we assume that $Q(\eta = \pm 1) = \frac{1}{2}$, then $c_n = \exp(\beta n)$, and the study [6] suggests that when performing a further random walk average

$$e^{-\beta n} Z_n^-(\beta) = E\left[e^{-\beta(H_n + n)}\right] = \exp\left(-\beta \chi n^{\frac{d}{d+2}} \log(n)^{\frac{2}{d+2}} (1 + o(1))\right). \tag{1.17}$$

and $\chi > 0$ is independent of β. Also, the proof of [6] suggests that, under the annealed measure, the walk is localized a time n into a ball of volume $(n/\log(n))^{\frac{d}{d+2}}$.

Our results focus on determining the correct temperature-scale, and are as follows.

Proposition 1.7 *Assume that $d = 3$, and $Q(\eta = \pm 1) = \frac{1}{2}$. The correct temperature-scaling is $1/n^{2/5}$. More precisely, there are positive constants $\beta_1 < \beta_2$, and the following holds. When $\beta > \beta_2$ (the low temperature regime), then for some positive constants a, c_1*

$$\exp(\beta n^{3/5}) \geq Z_n^- \left(\frac{\beta}{n^{2/5}}\right) \geq \exp(c_1 \beta n^{3/5}), \tag{1.18}$$

and,

$$\lim_{n\to\infty} P^-_{n,\frac{\beta}{n^{2/5}}} \left(|\{z \in \mathbb{Z}^d : \frac{n^{\frac{2}{5}}}{a} \leq l_n(z) \leq a n^{\frac{2}{5}}\}| \geq \frac{n^{3/5}}{a^4} \right) = 1. \tag{1.19}$$

When $\beta < \beta_1$ (the high temperature regime), for c_d defined in (1.2),

$$\lim_{n\to\infty} \frac{1}{n^{1/5}} \log Z_n^- \left(\frac{\beta}{n^{2/5}}\right) = \frac{c_d \beta^2}{2}. \tag{1.20}$$

Moreover, there is a positive constant b, such that

$$\lim_{n\to\infty} P^-_{n,\frac{\beta}{n^{2/5}}} \left(\{z \in \mathbb{Z}^d : l_n(z) \geq b n^{1/5}\} \neq \emptyset \right) = 0. \tag{1.21}$$

Remark 1.8 *We stress that (1.21) is not the 'correct' result, since we expect that in the high temperature regime, the polymer behaves like a random walk and we conjecture rather that for large b*

$$\lim_{n\to\infty} P^-_{n,\frac{\beta}{n^{2/5}}} \left(\{z \in \mathbb{Z}^d : \; l_n(z) \geq b \log(n)\} \neq \emptyset \right) = 0. \tag{1.22}$$

We include (1.21) to show the difference with (1.19) which occurs in the low temperature regime.

The paper is organized as follows. In Section 2, we recall the large deviations for the q-norm of the local times. We have then divided Theorems 1.1, 1.2, and 1.3, into their upper bounds parts, and their lower bounds parts. Upper bounds are treated in Section 3, while lower bounds are treated in Section 4. Finally, Section 5 contains the proof of Proposition 1.7.

2 Preliminaries

2.1 Sums of Independent variables

A. Nagaev has considered in [17] a sequence $\{\bar{Y}_n, n \in \mathbb{N}\}$ of independent centered i.i.d satisfying $\mathcal{H}_\alpha$ with $0 < \alpha < 1$, and has obtained the following upper bound (see also inequality (2.32) of S.Nagaev [18]).

Proposition 2.1 *Assume $E[\bar{Y}_i]=0$ and $E[(\bar{Y}_i)^2]\le 1$. There is a constant C_Y, such that for any integer n and any positive t*

$$P\left(\bar{Y}_1+\cdots+\bar{Y}_n\ge t\right)\le C_Y\left(nP\left(\bar{Y}_1>\frac{t}{2}\right)+\exp\left(-\frac{t^2}{20n}\right)\right). \tag{2.1}$$

Remark 2.2 *Note that if $\eta\in\mathcal{H}_\alpha$ for $1<\alpha\le 2$, then $\eta^2\in\mathcal{H}_{\frac{\alpha}{2}}$. Thus, for $\bar{Y}_i=\eta(i)^2-1$, Proposition 2.1 yields*

$$P\left(\sum_{i=1}^{n}(\eta(i)^2-1)\ge\xi_n\right)\le C_Y\left(n\exp\left(-c_\alpha(\xi_n)^{\alpha/2}\right)+\exp\left(-\frac{\xi^2 n^{2\beta-1}}{20}\right)\right). \tag{2.2}$$

Finally, we specialize to our setting a general lower bound of S.Nagaev (see Theorem 1 of [19]). Let $\{\Lambda_n,\ n\in\mathbb{N}\}$ a sequence of subsets of $\mathbb{Z}^d$, and for each n, let $\{Y_z^{(n)}, z\in\Lambda_n\}$ be independent and centered random variables. Let

$$\sigma_n^2=\sum_{z\in\Lambda_n}E\left[(Y_z^{(n)})^2\right],\quad\text{and}\quad \mathcal{C}_n^3=\sum_{z\in\Lambda_n}E\left[|Y_z^{(n)}|^3\right].$$

Proposition 2.3 *Consider a sequence $\{t_n,\ n\in\mathbb{N}\}$ such that for a small enough $\epsilon_{\mathcal{N}}>0$ and n large enough*

$$1\le t_n\le\ \epsilon_{\mathcal{N}}\ \min\Big(\frac{\sigma_n^3}{\mathcal{C}_n^3},\sigma_n(\max_{z\in\Lambda_n}\sqrt{E[(Y_z^{(n)})^2]})^{-1}\ \Big), \tag{2.3}$$

then, there is a positive constant κ such that

$$P\left(\frac{1}{\sigma_n}\sum_{z\in\Lambda_n}Y_z^{(n)}\ge t_n\right)\ge\exp\left(-\frac{t_n^2}{2}(1+\epsilon_{\mathcal{N}}\kappa)\right). \tag{2.4}$$

2.2 On self-intersection local times

In this section, we recall and establish useful estimates for functionals of the local times. First, for any $z\in\mathbb{Z}^d$, we estimate the variance of $q_n^2(z)-l_n(z)$

$$q_n^2(z)-l_n(z)=\big(\sum_{i\le l_n(z)}\eta_z(i)\big)^2-l_n(z)=\sum_{i\le l_n(z)}(\eta_z^2(i)-1)+2\sum_{1\le i<j\le l_n(z)}\eta_z(i)\eta_z(j), \tag{2.5}$$

It is immediate to obtain, for $\chi_1=E[\eta^4]+1$

$$2\left(l_n^2(z)-l_n(z)\right)\le E_Q\left[(q_n^2(z)-l_n(z))^2\right]=l_n(z)\left(E_Q[\eta^4]-1\right)+2\left(l_n^2(z)-l_n(z)\right)\le\chi_1 l_n^2(z). \tag{2.6}$$

Second, we summarize the asymptotic behavior of the q-norm of local times (for any real $q>1$)

$$\|l_n\|_q^q=\sum_{z\in\mathbb{Z}^d}l_n^q(z). \tag{2.7}$$

In dimension three and more, Becker and König [5] have shown that there are positive constants, say $\kappa(q,d)$, such that almost surely

$$\lim_{n\to\infty} \frac{\|l_n\|_q^q}{n} = \kappa(q,d). \tag{2.8}$$

The large deviations, and central limit theorem for $\|l_n\|_q$ are tackled in [2]: we establish a shape transition in the walk's strategy to realize the deviations $\{\|l_n\|_q^q - E[\|l_n\|_q^q] \ge n\xi\}$ with $\xi > 0$. This transition occurs at a critical value $q_c(d) = \frac{d}{d-2}$ suggesting the following picture.

- In the *super-critical regime* $q > q_c(d)$, the walk performs a short-time clumping on finitely many sites.
- In the *sub-critical regime* $q < q_c(d)$, the walk is localized during the whole time-period in a ball of volume $n/\xi^{\frac{1}{q-1}}$ where it visits each site of the order of $\xi^{\frac{1}{q-1}}$-times.

We first recall Theorem 1.2 of [2] which deals with the *super-critical* regime.

Lemma 2.4 *Assume $d \ge 3$ and $q > q_c(d)$. There are constants $C, c(q,d)$ (depending only on d and q), such that for $\xi_n \ge 1$, and any integer n*

$$\mathbb{P}_0\left(\|l_n\|_q^q - \mathbb{E}_0\left[\|l_n\|_q^q\right] > \xi_n n\right) \le C\exp\left(-c(q,d)(\xi_n\ n)^{\frac{1}{q}}\right). \tag{2.9}$$

Also, Lemma 1.4 of [2] estimates the cost of the contribution of *low* level sets to an excess q-norm. Thus, define for $x, y > 0$

$$\mathcal{D}_n(x,y) := \{z:\ x < l_n(z) \le y\}.$$

Lemma 2.5 *Assume $d \ge 3$ and $q \ge q_c(d)$. For $\gamma \ge 1$, and $\chi > 0$ and $\epsilon > 0$, there is a constant C such that for any sequence y_n*

$$\mathbb{P}_0\left(\sum_{z\in\mathcal{D}_n(1,y_n)} l_n^q(z) \ge \chi n^\gamma\right) \le C\exp\left(-\frac{n^{\gamma/q_c(d)-\epsilon}}{y_n^{(q/q_c(d)-1)}}\right). \tag{2.10}$$

When $\gamma = 1$, one needs to take $\chi > \kappa(q,d)$ in (2.10).

Remark 2.6 *Actually Lemma 1.4 of [2] is only stated for $\gamma > 1$. An inspection of its proof, shows that it covers also the case $\gamma = 1$ provided that $\chi > \kappa(q,d)$. In (2.10), we are unable to get rid of the ϵ. This is a delicate issue which is also responsible for a gap in the exponent of the speed in Region III of [4] (inequality (8)).*

The next result deals with *sub-critical regime.* It follows from Theorem 1.1 and Remark 1.3 of [2].

Lemma 2.7 *Assume $d \geq 3$ and $1 < q < q_c(d)$. There are constants $C, c(q,d)$ (depending only on d and q), such that for $\xi_n \geq 1$, and any integer n*

$$\mathbb{P}_0\left(\|l_n\|_q^q - \mathbb{E}_0\left[\|l_n\|_q^q\right] > \xi_n n\right) \leq C \exp\left(-c(q,d)\xi_n^{\frac{2}{d}\frac{1}{q-1}} n^{1-\frac{2}{d}}\right). \tag{2.11}$$

Remark 2.8 *For $d = 3$, (2.11) is mistakenly reported in [3]. Fortunately, this is of no consequence since (with the notations of [3] and in the so-called Region II), we need there*

$$\frac{2}{3}(\beta + b) - \frac{1}{3} - \epsilon > \beta - b \Longleftrightarrow 5\frac{\beta}{\alpha+1} > \beta + 1 + 3\epsilon \Longleftrightarrow \beta > \frac{\alpha+1}{4-\alpha}.$$

This latter condition defines Region II.

We now state a corollary of Lemmas 2.5 and 2.7, whose immediate proof is omitted.

Corollary 2.9 *Assume $d \geq 3$ and $\xi_n \geq n^{\frac{2}{3}}$. For $\epsilon > 0$ small enough, and n large enough*

$$\mathbb{P}_0\left(\|l_n\|_2 \geq \xi_n n^{-\epsilon}\right) \leq \exp\left(-\xi_n^{\frac{d}{d+2}} n^{\epsilon}\right). \tag{2.12}$$

3 Upper Bounds.

In this section, we prove the upper bounds in Theorems 1.1, 1.2, and 1.3. When dealing with large deviations, a natural approach is to perform a Chebychev's exponential inequality. If we expect $P(X_n \leq -x_n) \sim \exp(-\zeta_n)$, then for $\lambda > 0$, and $y_n = x_n/\zeta_n$

$$P\left(\langle l_n, 1 - \zeta_.(l_n)\rangle \geq x_n\right) \leq e^{-\lambda\zeta_n}\ E\left[\exp\left(\lambda\left\langle \frac{l_n}{y_n}, 1 - \zeta_.(l_n)\right\rangle\right)\right]. \tag{3.1}$$

Now, to get rid of the dependence between field and local time, we first perform an integration over the charges. We define for $x \in \mathbb{R}^+$ and $n \in \mathbb{N}$

$$\tilde{\Gamma}(x,n) = \log E_Q\left[\exp\left(x(1 - \zeta_0(n))\right)\right]. \tag{3.2}$$

Since $1 - \zeta_0(n) \leq 1$, and since $e^u \leq 1 + u + u^2$ when $u \leq 1$, we have, for the constant χ_1 which appears in (2.6),

$$\begin{aligned}\tilde{\Gamma}(x,n) &\leq \mathbb{1}_{\{x\geq 1\}}x + \mathbb{1}_{\{x<1\}}\ \log E_Q\left[1 + x(1-\zeta_0(n)) + x^2(1-\zeta_0(n))^2\right] \\ &\leq \mathbb{1}_{\{x\geq 1\}}x + \mathbb{1}_{\{x<1\}}\log\left(1 + x^2\mathrm{var}(\zeta_0(n))\right) \\ &\leq \mathbb{1}_{\{x\geq 1\}}x + \mathbb{1}_{\{x<1\}}x^2\sup_k \mathrm{var}(\zeta_0(k)) \leq \mathbb{1}_{\{x\geq 1\}}x + \mathbb{1}_{\{x<1\}}\ \chi_1 x^2.\end{aligned} \tag{3.3}$$

Remark 3.1 *Note first that (3.3) implies that $\tilde{\Gamma}(x,n) \leq \max(1,\chi_1)x^2$. Secondly, the dependence of $\tilde{\Gamma}(x,n)$ on the local times has vanished in these two regimes.*

Using (3.1) and (3.2), our first step is

$$P\left(\langle l_n, 1-\zeta_\cdot(l_n)\rangle \geq x_n\right) \leq e^{-\lambda\zeta_n}\ \mathbb{E}_0\left[\exp\left(\sum_{z\in\mathbb{Z}^d}\tilde{\Gamma}(\frac{\lambda l_n(z)}{y_n}, l_n(z))\right)\right]. \tag{3.4}$$

We introduce some notations. For $0<x<y$, and $\chi>0$

$$\mathcal{D}_n(x,y)=\left\{z\in\mathbb{Z}^d:\ x<l_n(z)\leq y\right\},\quad \text{and}\quad \mathcal{B}(x,y;\chi)=\left\{\sum_{z\in\mathcal{D}_n(x,y)} l_n^2(z)\geq\chi\right\}. \tag{3.5}$$

Also, we add a handy notations: for a subset $\Lambda\subset\mathbb{Z}^d$, $X_n(\Lambda)=\sum_{z\in\Lambda}X_n(z)$.

To treat separately the contribution of the two regimes of $\tilde{\Gamma}$, we divide the visited sites of the walk into $\mathcal{D}_n(1,y_n)$, and $\mathcal{D}_n(y_n,n)$. For $x_n'=x_n''=x_n/2$, and $0<\lambda<1$, we abbreviate $\mathcal{B}(1,y_n;\chi y_n x_n)$ by $\mathcal{B}$, and we have

$$\begin{aligned}
P\big(-X_n\geq x_n\big) \leq&\mathbb{P}_0\left(l_n(\mathcal{D}_n(y_n,n))\geq x_n'\right)+P\big(-X_n(\mathcal{D}_n(1,y_n))\geq x_n''\big)\\
\leq&\mathbb{P}_0\left(l_n(\mathcal{D}_n(y_n,n))\geq x_n'\right)+\mathbb{P}_0\left(\mathcal{B}\right)+P\big(-X_n(\mathcal{D}_n(1,y_n))\geq x_n'',\ \mathcal{B}^c\big)\\
\leq&\mathbb{P}_0\left(l_n(\mathcal{D}_n(y_n,n))\geq x_n'\right)+\mathbb{P}_0\left(\mathcal{B}\right)\\
&+\exp\left(-\lambda\frac{x_n''}{y_n}\right)\ \mathbb{E}_0\left[\ \mathbb{I}_{\mathcal{B}^c}\exp\left(\chi_1\lambda^2\sum_{\mathcal{D}_n(1,y_n)}\left(\frac{l_n(z)}{y_n}\right)^2\right)\right]\\
\leq&\mathbb{P}_0\left(l_n(\mathcal{D}_n(y_n,n))\geq x_n'\right)+\mathbb{P}_0\left(\mathcal{B}\right)+\exp\left(-\zeta_n(\frac{\lambda}{2}-\lambda^2\chi_1\chi)\right).
\end{aligned} \tag{3.6}$$

Note that the occurrence of an l_2-norm of the local time, in $\mathcal{B}(1,y_n;\chi)$, is not arbitrary but is a consequence of the asymptotic of the log-Laplace in (3.3).

We discuss now the respective contributions of the *top level term* $\{l_n(\mathcal{D}_n(y_n,n))\geq x_n'\}$, and of the *bottom level term* $\mathcal{B}(1,y_n;\chi y_n x_n)$. Note that the threshold y_n defining the *top level term* is determined by the log-Laplace, and may not be the value of the level set having a dominant contribution to our large deviation.

Top level term. First, note that for any $q>1$,

$$\{l_n(\mathcal{D}_n(y_n,n))\geq x_n'\}\subset\left\{\|\ \mathbb{I}_{\mathcal{D}_n(y_n,n)}l_n\|_q^q\geq\frac{1}{2}x_n y_n^{q-1}\right\}. \tag{3.7}$$

The event on the right hand side of (3.7) has a small probability if $x_n y_n^{q-1}>\kappa(q,d)n$, where $\kappa(q,d)$ is defined in (2.8).

We distinguish $q<q_c(d)$ and $q>q_c(d)$ with $q_c(d)=d/(d-2)$ (see Section 2.2). (i) When $q<q_c(d)$, the so-called *subcritical regime*, Lemma 2.7 yields

$$P\left(\|\ \mathbb{I}_{\mathcal{D}_n(y_n,n)}l_n\|_q^q\geq\frac{1}{2}x_n y_n^{q-1}\right)\leq\exp\left(-c(q,d)\left(\frac{x_n}{2n}y_n^{q-1}\right)^{\frac{2}{d}\frac{1}{(q-1)}}n^{1/q_c(d)}\right). \tag{3.8}$$

Now, since $x_n\leq n$, the map $q\mapsto\frac{x_n}{n}^{\frac{1}{(q-1)}}$ increases on $[1,q_c(d)[$. (ii) When $q>q_c(d)$, it is easy to check that the upper bound given by Lemma 2.5, increases on $]q_c(d),\infty[$, as a function of q. Thus, the best estimates we can obtain on $\{l_n(\mathcal{D}_n(y_n,n))\geq x_n'\}$ is with a bound as in (3.7) right at $q_c(d)$, for which we do not have sharp estimates.

Bottom level term. When $2 < q_c(d)$ (that is in $d = 3$), we expect $\mathcal{B}(1, y_n; \chi y_n x_n)$ to be of order $\{\|l_n\|_2^2 \geq \chi y_n x_n\}$, and by Lemma 2.7, we have in $d = 3$, for $\chi x_n y_n > \kappa(2, d)n$, that

$$P(\mathcal{B}(1, y_n; \chi y_n x_n)) \leq P\left(\|l_n\|_2^2 \geq \chi y_n x_n\right) \leq \exp\left(-c(2,3)(\chi y_n x_n)^{2/3} n^{-1/3}\right). \tag{3.9}$$

In this case, the cost of the bottom level set dominates the top level sets, and it is therefore useless to consider $q > 2$ in (3.8), when $d = 3$. When $q_c(d) \leq 2$ (that is when $d \geq 4$), and $x_n y_n / n \to \infty$, we can use Lemma 2.5, even though this is not an optimal result.

It is clear from this discussion that the behavior of the lower tail is distinct in $d = 3$ and in $d \geq 4$. This leads to different strategies, and different exponents. We discuss separately the case $d = 3$ and the case $d \geq 4$.

3.1 Dimension 3

We first make explicit the notations of (3.1)

$$x_n = \xi_n n^{\frac{2}{3}}, \quad \zeta_n = \xi_n^{\frac{4}{5}} n^{\frac{1}{3}}, \quad \text{and} \quad y_n = \frac{x_n}{\zeta_n} = \xi_n^{\frac{1}{5}} n^{\frac{1}{3}}. \tag{3.10}$$

where ξ_n can vary in $[a_0, n^{\frac{1}{3}}]$, for a constant a_0 to be specified later. Our first result is the following rough upper bound.

Lemma 3.2 *Assume $d = 3$. There are positive constants a_0, c_3^+, such that for $\xi_n \in [a_0, n^{\frac{1}{3}}]$*

$$P(-X_n \geq \xi_n n^{2/3}) \leq 3 \exp\left(-c_3^+ \xi_n^{\frac{4}{5}} n^{\frac{1}{3}}\right). \tag{3.11}$$

Note that in Section 4.2, we establish a similar lower bound.

Proof of Lemma 3.2 Recall that (3.7), for $q = 2$, requires that $x_n y_n > 2\kappa(2,3)n$, which is equivalent to $\xi_n > a_0 := (2\kappa(2,3))^{5/6}$. Recall that (3.9) requires that $\chi x_n y_n > \kappa(2,3)n$, which is equivalent to $\chi \xi_n^{6/5} > \kappa(2,3)$, which in turn requires that $\chi > 1/2$. Combining inequalities (3.6), (3.7) with $q = 2$, and (3.9), we obtain for $0 \leq \lambda \leq 1$

$$P(-X_n \geq \xi_n n^{2/3}) \leq \exp\left(-\frac{c(2,3)}{2^{2/3}}\zeta_n\right) + \exp\left(-c(2,3)\chi^{2/3}\zeta_n\right) + \exp\left(-(\frac{\lambda}{2} - \lambda^2 \chi_1 \chi)\zeta_n\right). \tag{3.12}$$

We choose $\chi = 1/4$, and $\lambda = \min(1/\chi_1, 1)$ in (3.9) to obtain the desired result.

∎

3.1.1 Upper bound in Theorem 1.1: $x_n = \xi_n n^{2/3} < n$

We show in this section that the dominant *level set* of the local times is of order $\xi_n^{\frac{6}{5}}$ much smaller than y_n when x_n is much smaller than n. We actually consider $x_n < a_1 n$ with a_1 to be chosen later small. For a large constant $a > 0$, to be chosen later, we decompose $\{z : l_n(z) > 0\}$ into $\mathcal{D}_1 \cup \cdots \cup \mathcal{D}_4$ with

$$\mathcal{D}_1 = \mathcal{D}_n(1, \frac{1}{a}\xi_n^{\frac{6}{5}}),\ \mathcal{D}_2 = \mathcal{D}_n(\frac{1}{a}\xi_n^{\frac{6}{5}}, a\xi_n^{\frac{6}{5}}),\ \mathcal{D}_3 = \mathcal{D}_n(a\xi_n^{\frac{6}{5}}, \frac{y_n}{a}),\ \text{and}\ \ \mathcal{D}_4 = \mathcal{D}_n(\frac{y_n}{a}, n). \tag{3.13}$$

We then write

$$P(-X_n \geq \xi_n n^{\frac{2}{3}}) \leq \sum_{i\neq 2} P\left(-X_n(\mathcal{D}_i) \geq \frac{1}{4}\xi_n n^{\frac{2}{3}}\right) + P\left(-X_n \geq \xi_n n^{\frac{2}{3}},\ -X_n(\mathcal{D}_2) \geq \frac{1}{4}\xi_n n^{\frac{2}{3}}\right). \tag{3.14}$$

We now show that the contribution of $\mathcal{D}_2$ is the dominant one.

a) Contribution of $\mathcal{D}_1$.

We use Chebychev's inequality with $\lambda > 0$,

$$P\left(-X_n(\mathcal{D}_1) \geq \frac{1}{4}\xi_n n^{2/3}\right) \leq e^{-\frac{\lambda}{4}\zeta_n}\mathbb{E}_0\left[\prod_{z\in\mathcal{D}_1}\exp\left(\tilde{\Gamma}(\frac{\lambda l_n(z)}{y_n}, l_n(z))\right)\right]. \tag{3.15}$$

Now, to justify the expansion of $\tilde{\Gamma}$ at 0, we need $\lambda\xi_n^{6/5} \leq a y_n$ which is equivalent to $\lambda\xi_n \leq a n^{1/3}$. Assume that this latter fact holds. We have by (3.3)

$$P\left(-X_n(\mathcal{D}_1) \geq \frac{1}{4}\xi_n n^{2/3}\right) \leq \exp\left(-\frac{\lambda}{4}\zeta_n + \chi_1\lambda^2\sum_{z\in\mathcal{D}_1}\frac{l_n^2(z)}{y_n^2}\right). \tag{3.16}$$

It will be convenient to define $\chi_2 = \max(\chi_1, \frac{1}{8})$. We now use that $l_n(\mathcal{D}_1) \leq n$, so that

$$\sum_{z\in\mathcal{D}_1}\frac{l_n^2(z)}{y_n^2} \leq \frac{\xi_n^{6/5}}{a y_n^2} l_n(\mathcal{D}_1) \leq \frac{\xi_n^{6/5} n}{a y_n^2} = \frac{\zeta_n}{a}. \tag{3.17}$$

We choose $\lambda = a/(8\chi_2) \leq a n^{1/3}/\xi_n$, and use (3.17) in (3.16)

$$P\left(-X_n(\mathcal{D}_1) \geq \frac{1}{4}\xi_n n^{2/3}\right) \leq \exp\left(-\frac{a}{8^2\chi_2}\zeta_n\right). \tag{3.18}$$

b) Contribution of $\mathcal{D}_3$.

For $0 \leq \lambda \leq a$, and χ to be chosen later, we have

$$\begin{aligned} P\left(-X_n(\mathcal{D}_3) \geq \frac{1}{4}\xi_n n^{2/3}\right) \leq & P\left(\mathcal{B}(a\xi_n^{6/5}, y_n; \chi x_n y_n)\right) + e^{-\frac{\lambda}{4}\zeta_n}\mathbb{E}_0\left[\ \mathbb{I}_{\mathcal{B}(.)^c}\exp\left(\chi_1\lambda^2\sum_{z\in\mathcal{D}_3}\frac{l_n^2(z)}{y_n^2}\right)\right] \\ \leq & P\left(\mathcal{B}(a\xi_n^{6/5}, y_n; \chi x_n y_n)\right) + \exp\left(-(\frac{\lambda}{4} - \chi_1\lambda^2\chi)\zeta_n\right). \end{aligned} \tag{3.19}$$

Choose $2 < q < q_c(3) = 3$, and by Lemma 2.7

$$\begin{aligned} P\left(\mathcal{B}(a\xi_n^{6/5}, y_n; \chi x_n y_n)\right) \leq & P\left(\|l_n\|_q^q \geq (a\xi_n^{6/5})^{q-2}\chi x_n y_n\right) = P\left(\|l_n\|_q^q \geq a^{q-2}\xi_n^{6/5(q-1)}\chi n\right) \\ \leq & \exp\left(-c(q,3)\left(a^{q-2}\chi\xi_n^{\frac{6}{5}(q-1)}\right)^{\frac{2}{3(q-1)}} n^{1/3}\right) \\ \leq & \exp\left(-c(q,3)\left(a^{q-2}\chi\right)^{\frac{2}{3(q-1)}}\zeta_n\right) \end{aligned} \tag{3.20}$$

Now, collecting (3.19) and (3.20), we choose $\chi = a^{1-q/2}$ and for $a^{4-q} > (8\chi_1)^{-2}$ we have that the optimal λ in (3.19) satisfies $\lambda \le a$, and

$$P\left(-X_n(\mathcal{D}_3) \ge \frac{1}{4}\xi_n n^{2/3}\right) \le \exp\left(-c(q,3)\left(a^{q-2}\chi\right)^{\frac{2}{3(q-1)}}\zeta_n\right) + \exp\left(-(\frac{\lambda}{4} - \chi_1\lambda^2\chi)\zeta_n\right)$$
$$\le \exp\left(-c(q,3)a^{\frac{q-2}{3(q-1)}}\zeta_n\right) + \exp\left(-\frac{1}{8^2\chi_1}a^{q/2-1}\zeta_n\right) \tag{3.21}$$

c) Contribution of $\mathcal{D}_4$.

We proceed as in (3.7) and (3.8).

$$P\left(-X_n(\mathcal{D}_4) \ge \frac{1}{4}\xi_n n^{2/3}\right) \le P\left(l_n(\mathcal{D}_4) \ge \frac{1}{4}\xi_n n^{2/3}\right) \le P\left(\|l_n\|_q^q \ge \frac{1}{4}\xi_n(\frac{y_n}{a})^{q-1}n^{2/3}\right)$$
$$\le \exp\left(-c(q,3)\left(\frac{\xi_n}{4n^{1/3}}(\frac{y_n}{a})^{q-1}\right)^{\frac{2}{3(q-1)}} n^{1/3}\right). \tag{3.22}$$

Now, for $A > 0$, and $2 < q < 3$,

$$\frac{1}{a^{2/3}}(\xi_n y_n^{q-1})^{\frac{2}{3(q-1)}}\ n^{\frac{1}{3}(1-\frac{2}{3(q-1)})} \ge A\xi_n^{4/5}n^{1/3} \quad \Longleftrightarrow \quad \xi_n(aA^{3/2})^{\frac{(q-1)}{(q-2)}} \le n^{1/3}. \tag{3.23}$$

Our assumption is that $\xi_n < a_1 n^{1/3}$, and this implies that

$$P\left(-X_n(\mathcal{D}_4) \ge \frac{1}{4}\xi_n n^{2/3}\right) \le \exp\left(-c(q,3)\frac{\zeta_n}{a_1^\gamma a^{2/3}}\right), \quad \text{with} \quad \gamma = \frac{2(q-2)}{3(q-1)} > 0. \tag{3.24}$$

d) Contribution of $\mathcal{D}_2$.

We recall the rough lower bound $P(-X_n \ge \xi_n n^{\frac{2}{3}}) \ge \exp\left(-c_3^-\zeta_n\right)$, and express (3.14) as

$$P(-X_n \ge \xi_n n^{\frac{2}{3}}) \le \sum_{i\neq 2} P\left(-X_n(\mathcal{D}_i) \ge \frac{1}{4}\xi_n n^{\frac{2}{3}}\right) + P\left(-X_n \ge \xi_n n^{\frac{2}{3}}, -X_n(\mathcal{D}_2) \ge \frac{1}{4}\xi_n n^{\frac{2}{3}}\right). \tag{3.25}$$

When a is large enough in (3.18) and (3.21), and a_1 small enough in (3.24), the terms with $\mathcal{D}_1$ and $\mathcal{D}_3$ are negligible. We then write

$$\left\{-X_n(\mathcal{D}_2) \ge \frac{1}{4}\xi_n n^{\frac{2}{3}}\right\} \subset \left\{|\mathcal{D}_2| \ge \frac{n}{a^4\xi_n^{6/5}}\right\} \cup \left\{\sum_{\mathcal{D}_2}(1-\zeta_z(l_n(z))) \ge \frac{n^{\frac{2}{3}}}{4a\xi_n^{1/5}},\ |\mathcal{D}_2| \le \frac{n}{a^4\xi_n^{6/5}}\right\}. \tag{3.26}$$

Now, for dealing with the last event in (3.26), note that

$$\left\{\sum_{\mathcal{D}_2}(1-\zeta_z(l_n(z))) \ge \frac{n^{\frac{2}{3}}}{4a\xi_n^{1/5}},\ |\mathcal{D}_2| \le \frac{n}{a^4\xi_n^{6/5}}\right\} \subset \left\{\frac{1}{\sqrt{|\mathcal{D}_2|}}\sum_{\mathcal{D}_2}(1-\zeta_z(l_n(z))) \ge \frac{a\xi_n^{2/5}n^{1/6}}{4}\right\}. \tag{3.27}$$

Now, we fix the randomness of the walk, and use that $1-\zeta_z \le 1$, $E_Q[1-\zeta_z]=0$ and $E_Q[(1-\zeta_z)^2] \le \chi_1$ to obtain that (recall that $\zeta_n = \xi_n^{4/5} n^{1/3}$)

$$P\left(\frac{1}{\sqrt{|\mathcal{D}_2|}} \sum_{\mathcal{D}_2} (1-\zeta_z(l_n(z))) \ge \frac{a\xi_n^{2/5} n^{1/6}}{4}\right) \le \exp(-\frac{a^2 \zeta_n}{4}). \tag{3.28}$$

We put together (3.25), (3.26) and (3.28) to obtain for a large enough

$$\lim_{n\to\infty} P\left(|\mathcal{D}_2| \ge \frac{n}{a^4 \xi_n^{6/5}} \;\Big\|\; -X_n \ge \xi_n n^{\frac{2}{3}}\right) = 1. \tag{3.29}$$

3.1.2 Upper bound in Theorem 1.1: $x_n = \xi n$ with $1 > \xi > a_1$.

Note that

$$\xi_n = \xi n^{1/3}, \quad \zeta_n = \xi^{4/5} n^{3/5}, \quad \text{and} \quad y_n = \xi^{1/5} n^{2/5}.$$

Note that $\xi_n^{6/5} = \xi y_n < y_n$. For a large constant $b > 0$, to be chosen later, we decompose $\{z : l_n(z) > 0\}$ into $\mathcal{D}_1 \cup \cdots \cup \mathcal{D}_3$ with

$$\mathcal{D}_1 = \mathcal{D}_n(1, \frac{1}{b}\xi^{\frac{6}{5}} n^{2/5}),\ \mathcal{D}_2 = \mathcal{D}_n(\frac{1}{b}\xi^{\frac{6}{5}} n^{2/5}, b\xi^{\frac{1}{5}} n^{2/5}),\ \text{and} \quad \mathcal{D}_3 = \mathcal{D}_n(by_n, n). \tag{3.30}$$

We then write

$$P(-X_n \ge \xi n) \le \sum_{i\neq 2} P\left(-X_n(\mathcal{D}_i) \ge \frac{1}{4}\xi n\right) + P\left(-X_n(\mathcal{D}_2) \ge \frac{1}{2}\xi n,\ -X_n \ge \xi n\right), \tag{3.31}$$

and we show that the contribution of $\mathcal{D}_2$ is the dominant one.

The treatment of $\mathcal{D}_1$ is similar to the previous case a). The choice $\lambda = b/(8\chi_2)$ requires $\xi \le 8\chi_2$, which holds since $\xi < 1 \le 8\chi_2$.

Then, for $\mathcal{D}_3$, we write

$$\begin{aligned} P\left(-X_n(\mathcal{D}_3) \ge \frac{1}{4}\xi n\right) \le & P\left(l_n(\mathcal{D}_3) \ge \frac{1}{4}\xi n\right) \le P\left(\|l_n\|_2^2 \ge \frac{1}{4} b\xi^{6/5} n^{2/5} n\right) \\ \le & \exp\left(-c(2,3)\left(\frac{b}{4}\right)^{2/3} \zeta_n\right). \end{aligned} \tag{3.32}$$

By taking b large enough, and proceeding as in the previous case d), we reach that for $\xi < 1$

$$\lim_{n\to\infty} P\left(|\mathcal{D}_2| \ge \frac{\xi^{4/5} n^{3/5}}{b} \;\Big\|\; X_n \le -\xi n\right) = 1. \tag{3.33}$$

3.2 Dimension 4 or more.

We choose here x_n, y_n and ζ_n as follows.

$$x_n = \xi_n \sqrt{n}, \quad \zeta_n = \xi_n^2, \quad \text{and} \quad y_n = \frac{\sqrt{n}}{\xi_n}. \tag{3.34}$$

We first deal with the case $a_0 n^{1/6} \le \xi_n \ll n^{\gamma_d - \epsilon}$, with $\gamma_d = (d/2)/(d+4)$, and any ϵ positive.

3.2.1 Proof of the Upper bound in (1.11).

Our starting point is the inequality (3.6) with x_n, y_n, ζ_n as in (3.34). We deal with each term on the right hand side of (3.6).

First, choose $\chi > \kappa(2,d)$, and Lemma 2.5 gives

$$P(\mathcal{B}(1, y_n; \chi x_n y_n)) = \mathbb{P}_0\left(\sum_{z\in\mathcal{D}_n(1,y_n)} l_n^2(z) \geq \chi n\right) \leq \exp\left(-\frac{n^{1/q_c(d)-\epsilon}}{y_n^{(2/q_c(d)-1)}}\right). \tag{3.35}$$

Second, $n^{1/q_c(d)-\epsilon} \geq y_n^{(2/q_c(d)-1)}\xi_n^2$ is equivalent to asking $\xi_n^{1+4/d} \leq n^{1/2-\epsilon}$, which is exactly the condition which defines this regime.

Now, we deal with the event $\{l_n(\mathcal{D}_n(y_n, n)) \geq x_n/2\}$. The proof of Proposition 3.3 of [4] yields

$$P\left(l_n(\mathcal{D}_n(y_n, n)) \geq x_n/2\right) \leq \exp\left(-x_n^{1/q_c(d)} y_n^{2/d}\right), \tag{3.36}$$

provided that for some fixed a and n large

$$y_n^{1+\frac{2}{d}} \geq \log^a(n) x_n^{2/d}. \tag{3.37}$$

Now both $x_n^{1/q_c(d)} y_n^{2/d} \gg \xi_n^2$ and condition (3.37) follow from $\log \xi_n \leq (d/2-\epsilon)/(d+4)\log(n)$. Thus, for any $\epsilon > 0$, there is $\epsilon' > 0$ such that

$$P\left(l_n(\mathcal{D}_n(y_n, n)) \geq x_n/2\right) \leq \exp\left(-n^{\epsilon'}\xi_n^2\right). \tag{3.38}$$

A bound of the type $P(-X_n \geq x_n) \leq \exp(-c\xi_n^2)$ now follows from (3.35), and (3.36) after we choose λ small enough in the last term of the right hand side of (3.6).

3.2.2 Proof of (1.12)

We fix A large constant, and take the subdivision $\{b_1, \ldots, b_M\}$ of $[A, y_n[$ with $b_1 = A$, $b_{i+1} = 2b_i$, for $i = 1, \ldots, M-1$, with M of order $\log(n)$. We will choose q slightly larger than 2, to be in the super-critical regime (when $d \geq 4$), and we define

$$\mathcal{G}_i = \left\{|\mathcal{D}_n(b_i, b_{i+1})| < \frac{C_1 n}{b_{i+1}^q}\right\}. \tag{3.39}$$

Finally, for $q > 2$, choose $p_i = p2^{-i(q-2)/2}$ where p is such that $\sum_i p_i = 1$. Now,

$$\begin{aligned} P\Big(\sum_i \sum_{z\in\mathcal{D}_n(b_i,b_{i+1})} l_n(z)(1-\zeta_z(l_n(z))) \geq x_n\Big) &\leq P\left(\cup_i \mathcal{G}_i^c\right) \\ &+ \sum_i P\left(\sum_{z\in\mathcal{D}_n(b_i,b_{i+1})} \frac{l_n(z)}{b_{i+1}}(1-\zeta_z(l_n(z))) \geq \frac{x_n}{b_{i+1}},\ \mathcal{G}_i\right). \end{aligned} \tag{3.40}$$

First, we deal with $P(\cup_i \mathcal{G}_i^c)$ in the right hand side of (3.40). Note that

$$\cup_i\, \mathcal{G}_i^c \subset \left\{\|\ \mathbb{I}_{\mathcal{D}_n(A,y_n)} l_n\|_q^q \geq \frac{C_1}{2^q} n\right\}. \tag{3.41}$$

We choose $C_1 = 2^{q+1}\kappa(q,d)$, and use Lemma 2.5 to obtain, for any $\epsilon' > 0$,

$$P\left(\cup_i \mathcal{G}_i^c\right) \le \exp\left(-\frac{n^{1/q_c(d)-\epsilon'}}{y_n^{q/q_c(d)-1}}\right). \tag{3.42}$$

We neglect $P(\cup \mathcal{G}_i^c)$ if $n^{1/q_c(d)-\epsilon'} \ge y_n^{q/q_c(d)-1}\xi_n^2$. Since $\log(\xi_n) \le (d/2-\epsilon)/(d+4)\log(n)$, and we are interested in q close to 2, we only need to check that taking $q=2$, for any $\epsilon > 0$, we can find $\epsilon' > 0$ such that

$$\frac{1}{q_c(d)} - \frac{1}{2}\left(\frac{2}{q_c(d)} - 1\right) - \epsilon' \ge \left(2 - (\frac{2}{q_c(d)} - 1)\right)\left(\frac{d/2-\epsilon}{d+4}\right) \Longleftrightarrow \frac{1}{2} - \epsilon' \ge \frac{1}{2} - \frac{\epsilon}{d}. \tag{3.43}$$

Since (3.43) holds, we can find $\delta > 0$ small enough, and $q = 2+\delta$ so that $P(\cup \mathcal{G}_i^c)$ is negligible.

We fix a realization of the random walk and integrate first with respect to charges. For the charges, we use the gaussian bounds of Remark 3.1 which states that $\tilde{\Gamma}(x,n) \le \bar{\chi}_1 x^2$, where $\bar{\chi}_1 = \max(1,\chi_1)$. In other words, on the event $\mathcal{G}_i = \{|\mathcal{D}_n(b_i,b_{i+1})| \le C_1 n/b_{i+1}^q\}$, we use

$$\begin{aligned} &Q\left(\sum_{i=1}^{M}\sum_{z\in\mathcal{D}_n(b_i,b_{i+1})} l_n(z)\left(1-\zeta_z(l_n(z))\right) > \sum_i p_i x_n\right) \\ &\qquad\qquad \le \sum_{i=1}^{M} Q\left(\sum_{z\in\mathcal{D}_n(b_i,b_{i+1})} \frac{l_n(z)}{b_{i+1}}\left(1-\zeta_z(l_n(z))\right) > \frac{p_i}{b_{i+1}} x_n\right). \end{aligned} \tag{3.44}$$

Now, we consider a fixed $i \in \{1,\dots,M\}$, and on $\mathcal{G}_i$, we have for any $\theta > 0$

$$\begin{aligned} Q\left(\sum_{z\in\mathcal{D}_n(b_i,b_{i+1})} \frac{l_n(z)}{b_{i+1}}\left(1-\zeta_z(l_n(z))\right) > \frac{p_i}{b_{i+1}} x_n\right) &\le \exp\left(-\frac{p_i x_n \theta}{b_{i+1}} + \bar{\chi}_1|\mathcal{D}_n(b_i,b_{i+1})|\theta^2\right) \\ &\le \exp\left(-\frac{p_i x_n \theta}{b_{i+1}} + \bar{\chi}_1 C_1 \frac{n}{b_{i+1}^q}\theta^2\right). \end{aligned} \tag{3.45}$$

Note that if $|\mathcal{D}_n(b_i,b_{i+1})| \le p_i x_n/b_{i+1}$, then the left hand side of (3.45) vanishes. Therefore, we assume that $|\mathcal{D}_n(b_i,b_{i+1})| > p_i x_n/b_{i+1}$, so that the θ which minimizes the right hand side of (3.45) is lower than 1, and we obtain

$$P\left(\sum_{z\in\mathcal{D}_n(b_i,b_{i+1})} \frac{l_n(z)}{b_{i+1}}\left(1-\zeta_z(l_n(z))\right) > \frac{p_i}{b_{i+1}} x_n,\ \mathcal{G}_i\right) \le \exp\left(-\frac{p_i^2 b_{i+1}^{q-2}\xi_n^2}{4C_1}\right). \tag{3.46}$$

With our choice of p_i, b_i, we have that $p_i^2 b_{i+1}^{q-2} \ge p^2 A^{q-2}$. Combining (3.44) and (3.46), we have

$$P(\sum_{z\in\mathbb{Z}^d} l_n(z) 1 - \zeta_z(l_n(z)) \ge x_n/2) \le M\exp\left(-\frac{p^2 A^{q-2}\xi_n^2}{4C_1}\right). \tag{3.47}$$

The bound (1.12) follows from (3.38) and (3.47).

3.2.3 Dimension $d \geq 4$, and $\frac{d+2}{d+4} < \beta < 1$.

This corresponds to Region III of [4]. We set $x_n = \xi_n$, $\zeta_n = \xi_n^{\frac{d}{d+2}}$, and $y_n = \xi_n/\zeta_n$. Instead of (3.6), we use

$$
\begin{aligned}
P\big(-X_n \geq \xi_n\big) \leq \mathbb{P}_0\left(l_n(\mathcal{D}_n(y_n^{1+\epsilon}, n)) \geq \frac{\xi_n}{2}\right) \\
+ \mathbb{P}_0\left(\| \mathbb{I}_{\mathcal{D}_n(1, y_n^{1+\epsilon})} l_n\|_2^2 \geq y_n \xi_n\right) + \exp\left(-\zeta_n y_n^{-\epsilon}(\lambda \xi_2 - \lambda^2 \chi_1)\right).
\end{aligned}
\tag{3.48}
$$

Proposition 3.3 of [4] yields that there is $\epsilon' > 0$ such that

$$
\mathbb{P}_0\left(l_n(\mathcal{D}_n(y_n^{1+\epsilon}, n)) \geq \frac{\xi_n}{2}\right) \leq \exp(-\xi_n^{\frac{d}{d+2}-\epsilon'}). \tag{3.49}
$$

Now $\zeta_n^{\frac{d+4}{d+2}} \geq n$, and by Lemma 2.5, for any ϵ

$$
\mathbb{P}_0\left(\sum_{z \in \mathcal{D}_n(1, y_n^{1+\epsilon})} l_n^2(z) \geq \xi_n^{\frac{d+4}{d+2}}\right) \leq \exp\Big(-\frac{\xi_n^{(\frac{d+4}{d+2})(\frac{1}{q_c(d)}-\epsilon)}}{y_n(\frac{2}{q_c(d)}-1)}\Big). \tag{3.50}
$$

The upper bound in (1.13) follows from (3.48), (3.50), and (3.49).

4 Lower Bounds.

In realizing the lower bounds for Theorems 1.1, 1.2, and 1.3, two strategies of the walk are distinguished: (i) the walk is localized a time T_n into a ball of radius r_n with $r_n^2 \ll T_n$, (ii) the walk roams freely.

4.1 On localizing the walk

We introduce two sequences $\{T_n, r_n,\ n \in \mathbb{N}\}$. We force the random walk to spend a time T_n in the ball centered at 0, of radius r_n, that we denote $B(r_n)$.

If $\tau_n = \inf\{n \geq 0 : S(n) \notin B(r_n)\}$, it is well known that for some constant c_0

$$
\mathbb{P}_0(\tau_n > T_n) \geq \exp\left(-c_0 \frac{T_n}{|B(r_n)|^{2/d}}\right). \tag{4.1}
$$

Once the walk is forced to stay inside $B(r_n)$, we turn to estimating the cost of $\{X_n < -x_n\}$. We then choose $\{T_n, r_n\}$ so as to match the cost with (4.1).

First, we need some relation between being localized a time T_n in a ball $B(r_n)$, and visiting enough sites of $B(r_n)$ a time of order $T_n/|B(r_n)|$. We have shown in [3] Proposition 1.4, that in $d = 3$, for sequences $\{r_n, T_n\}$ going to infinity with $r_n^d \leq K T_n$, for some constant K, there are positive constants δ_0 and ϵ_0, independent of r_n, T_n such that, for n large enough

$$
\mathbb{P}_0\left(|\{z :\ l_{T_n}(z) > \delta_0 \frac{T_n}{|B(r_n)|}\}| \geq \epsilon_0 |B(r_n)|\right) \geq \frac{1}{2}\mathbb{P}_0\left(\tau_n > T_n\right). \tag{4.2}
$$

Let $\mathcal{R}_n$ be the set of sites visited by the random walk before time n. The only fact used in proving (4.2) is an asymptotical bound on $\mathbb{P}_0(|\mathcal{R}_n| < n/\xi)$ for a fixed large ξ and n going to infinity. Now, there is an obvious relation between $|\mathcal{R}_n|$ and $\|l_n\|_q$ which reads as follows. For $q > 1$

$$\left(\frac{n}{|\mathcal{R}_n|}\right)^{q-1} \leq \frac{\|l_n\|_q^q}{n}. \tag{4.3}$$

Thus, from (4.3) and [2] Theorem 1.1, we have for $\xi^{q-1} > \kappa(q,d)$, and $q < q_c(d)$

$$\mathbb{P}_0\left(|\mathcal{R}_n| < \frac{n}{\xi}\right) \leq \mathbb{P}_0\left(\|l_n\|_q^q \geq \xi^{q-1} n\right) \leq \exp(-c_1^+ \xi^{\frac{2}{d}} n^{1-\frac{2}{d}}). \tag{4.4}$$

Since $q_c(d) = \frac{d}{d-2} > 1$, as soon as $d \geq 3$, (4.4) is sufficient to obtain (4.2) by following the proof of [3], and we omit the details. We now focus on the following set of sites

$$\mathcal{G}_n = \left\{ z :\ \delta_0 \frac{T_n}{|B(r_n)|} \leq l_{T_n}(z) \leq \frac{2T_n}{\epsilon_0 |B(r_n)|} \right\}. \tag{4.5}$$

Note that

$$|\{z :\ l_{T_n}(z) > \frac{2T_n}{\epsilon_0|B(r_n)|}\}| \leq \frac{\epsilon_0}{2}|B(r_n)|,$$

so that $\{l_{T_n} > \delta_0 T_n/|B(r_n)|\} = \mathcal{G}_n \cup \{l_{T_n} > 2T_n/(\epsilon_0|B(r_n)|)\}$, and

$$\mathbb{P}_0\left(|\mathcal{G}_n| \geq \frac{\epsilon_0}{2}|B(r_n)|\right) \geq \mathbb{P}_0\left(|\{z :\ l_{T_n}(z) > \delta_0 \frac{T_n}{|B(r_n)|}\}| \geq \epsilon_0 |B(r_n)|\right). \tag{4.6}$$

Now, in the scenario we are adopting, it will be easy to estimate the contribution of sites of $\mathcal{G}_n$, which is a random set. To use the notations of Proposition 2.3, we define for $z \in \mathbb{Z}^d$, $Y_z^{(n)} = l_n(z)(1 - \zeta_z(l_n(z)))$. We have, for $\delta > 0$ small

$$\left\{ \sum_{z \in \mathbb{Z}^d} Y_z^{(n)} \geq x_n \right\} \supset \left\{ \sum_{z \in \mathcal{G}_n} Y_z^{(n)} \geq (1+\delta) x_n \right\} \cap \left\{ \sum_{z \notin \mathcal{G}_n} Y_z^{(n)} \geq -\delta x_n \right\}. \tag{4.7}$$

When we integrate (4.7) over the charges, we use that charges over disjoint regions are independent. Thus, we fix a realization of the walk, and

$$Q\left(\sum_{z \in \mathbb{Z}^d} Y_z^{(n)} \geq x_n \right) \geq Q\left(\sum_{z \in \mathcal{G}_n} Y_z^{(n)} \geq (1+\delta) x_n \right) Q\left(\sum_{z \notin \mathcal{G}_n} Y_z^{(n)} \geq -\delta x_n \right). \tag{4.8}$$

We first deal with the charges in $\mathcal{G}_n^c$. We show using (2.6) that on $\mathcal{B}_n = \{\|l_n\|_2 \leq x_n n^{-\epsilon'}\}$, for ϵ' small, then

$$\begin{aligned} \mathbb{1}_{\mathcal{B}_n} Q\left(\sum_{z \notin \mathcal{G}_n} Y_z^{(n)} \leq -\delta x_n \right) &\leq \ \mathbb{1}_{\mathcal{B}_n} \frac{\sum_{z \in \mathbb{Z}^d} E[(Y_z^{(n)})^2]}{(\delta x_n)^2} \\ &\leq \ \mathbb{1}_{\mathcal{B}_n} \frac{\chi_1 \sum_{z \in \mathbb{Z}^d} l_n^2(z)}{(\delta x_n)^2} \leq \ \mathbb{1}_{\mathcal{B}_n} \frac{\chi_1}{\delta^2 n^{2\epsilon'}}. \end{aligned} \tag{4.9}$$

Thus, from (4.9) with n large, we have

$$\mathbb{I}_{\mathcal{B}_n} Q\left(\sum_{z\notin\mathcal{G}_n} Y_z^{(n)} \geq -\delta x_n\right) \geq \frac{\mathbb{I}_{\mathcal{B}_n}}{2} \tag{4.10}$$

From (4.7) and (4.10), we obtain, when integrating only over the charges

$$\mathbb{I}_{\mathcal{B}_n} Q\left(\sum_{z\in\mathbb{Z}^d} Y_z^{(n)} \geq x_n\right) \geq \frac{\mathbb{I}_{\mathcal{B}_n}}{2} Q\left(\sum_{z\in\mathcal{G}_n} Y_z^{(n)} \geq (1+\delta)x_n\right). \tag{4.11}$$

Thus, after integrating over the walk

$$\begin{aligned} 2P\left(\sum_{z\in\mathbb{Z}^d} Y_z^{(n)} \geq x_n\right) + \mathbb{P}_0\left(\mathcal{B}_n^c\right) &\geq P\left(\sum_{z\in\mathcal{G}_n} Y_z^{(n)} \geq (1+\delta)x_n\right) \\ &\geq P\left(|\mathcal{G}_n| \geq \frac{\epsilon_0}{2}|B(r_n)|, \sum_{z\in\mathcal{G}_n} Y_z^{(n)} \geq (1+\delta)x_n\right). \end{aligned} \tag{4.12}$$

Assume for a moment that $\mathbb{P}_0(\mathcal{B}_n^c)$ were negligible. When integrating only over charges the last term of (4.12), we invoke Nagaev's Proposition 2.3, applied to $\{Y_z^{(n)},\ z \in \mathcal{G}_n\}$. To simplify notations, we assume henceforth that $T_n = n$ (though we can force the transient walk never to return to $\mathcal{G}_n$ after time T_n, so that for $z \in \mathcal{G}_n$ we would have $l_n(z) = l_{T_n}(z)$). Now, when we fix a realization of the walk, we have easily from the equality (2.6), for constants χ_1 and χ_4

$$\chi_1 l_n^2(z) \geq E_Q[(Y_z^{(n)})^2] \geq 2(l_n^2(z) - l_n(z)) \quad \text{and} \quad E_Q[(Y_z^{(n)})^4] \leq \chi_4 l_n^4(z). \tag{4.13}$$

From Jensen's inequality, we have $E_Q[|Y_z^{(n)}|^3] \leq \chi_3 l_n^3(z)$ with $\xi_3 = \xi_4^{3/4}$. Note that in order to have a non-zero lower bound for the variance of $Y_z^{(n)}$, we impose

$$\delta_0 \frac{T_n}{|B(r_n)|} \geq 2 \quad \text{so that} \quad \forall z \in \mathcal{G}_n \quad E_Q[Y_z^2] \geq 2(l_n^2(z) - l_n(z)) \geq l_n^2(z). \tag{4.14}$$

With the notations of Proposition 2.3, we have (using (4.13)) on $\{|\mathcal{G}_n| \geq \frac{\epsilon_0}{2}|B(r_n)|\}$

$$\frac{\epsilon_0\delta_0^2}{2}\frac{T_n^2}{|B(r_n)|} \leq \sigma_n^2 \leq \frac{4\chi_1}{\delta_0\epsilon_0^2}\frac{T_n^2}{|B(r_n)|} \quad \text{and} \quad \mathcal{C}_n^3 \leq \frac{8\chi_3}{\delta_0\epsilon_0^3}\frac{T_n^3}{|B(r_n)|^2}. \tag{4.15}$$

Also, $\sigma_n t_n = (1+\delta)x_n$, so that (2.3) holds if for some $\epsilon_{\mathcal{N}} > 0$, and n large enough

$$\sigma_n \leq (1+\delta)x_n, \quad (1+\delta)x_n\mathcal{C}_n^3 \leq \epsilon_{\mathcal{N}}\sigma_n^4, \quad \text{and} \quad (1+\delta)x_n \max_{z\in\mathcal{G}_n}\sqrt{E\left[(Y_z^{(n)})^2\right]} \leq \epsilon_{\mathcal{N}}\sigma_n^2. \tag{4.16}$$

Using (4.15), (4.16) and (4.14) follow if, for some constant c_1

$$\frac{4\chi_1}{\delta_0\epsilon_0^2}\frac{T_n^2}{|B(r_n)|} \leq x_n^2, \quad \text{and} \quad x_n \leq \epsilon_{\mathcal{N}} c_1 T_n. \tag{4.17}$$

When (4.17) holds, and we can use Proposition 2.3, to obtain on $\{|\mathcal{G}_n| \geq \frac{\epsilon_0}{2}|B(r_n)|\}$, and for constants c_1, c_2

$$Q\left(\sum_{z\in\mathcal{G}_n} Y_z^{(n)} \geq (1+\delta)x_n\right) \geq \exp\left(-c_1(\frac{x_n}{\sigma_n})^2\right) \geq \exp\left(-c_2\frac{x_n^2\ |B(r_n)|}{T_n^2}\right). \tag{4.18}$$

After integrating over the walk, recalling (4.2), (4.1), (4.12) and (4.6), we have

$$\begin{aligned}2P\left(\sum_{z\in\mathbb{Z}^d} Y_z^{(n)} \geq x_n\right) &\geq P\left(|\mathcal{G}_n| \geq \frac{\epsilon_0}{2}|B(r_n)|, \sum_{z\in\mathcal{G}_n} Y_z^{(n)} \geq (1+\delta)x_n\right) - \mathbb{P}_0(\mathcal{B}_n^c)\\ &\geq \exp\left(-c_2\frac{x_n^2\ |B(r_n)|}{T_n^2} - c_0\frac{T_n}{|B(r_n)|^{2/d}}\right) - \mathbb{P}_0(||l_n||_2^2 \geq x_n^2 n^{-2\epsilon'}).\end{aligned} \tag{4.19}$$

From inequality (4.19), the difference between $d=3$ and $d\geq 4$ is obvious, when imposing a localisation of the walk. Indeed, matching the two costs in (4.19), we find

$$\frac{x_n^2|B(r_n)|}{T_n^2} = \frac{T_n}{|B(r_n)|^{2/d}} \Longrightarrow |B(r_n)|^{\frac{d+2}{d}} = \frac{T_n^3}{x_n^2}. \tag{4.20}$$

Thus, combining (4.19) with the choice of (4.20), we obtain for a constant $c_d^- > 0$

$$P(X_n \leq -x_n) \geq \exp\left(-c_d^- x_n^{\frac{4}{d+2}} T_n^{\frac{d-4}{d+2}}\right) - \mathbb{P}_0(\mathcal{B}_n^c). \tag{4.21}$$

Corollary 2.9 shows that $\mathbb{P}_0(\mathcal{B}_n^c) \ll \exp(-c_d^- \xi_n^{\frac{d}{d+2}})$. Henceforth, we neglect $\mathbb{P}_0(\mathcal{B}_n^c)$.

4.2 The case $d=3$ and $a_0 \leq \xi_n \leq n^{1/3}$.

In this section, we choose $T_n = n$, and $|B(r_n)|^{5/3} = n^3/x_n^2$, as suggested in (4.20).

We start with $\xi_n \leq c_1\epsilon_{\mathcal{N}} n^{1/3}$. In this case, $x_n = \xi_n n^{2/3}$. The discussion of the previous section applies here. Note that sites of $\mathcal{G}_n$ are visited about $\xi_n^{6/5}$-times each. Conditions (4.17) are satisfied, and the discussion following it holds. The bound (4.21) provides the desired lower bound.

Now, we deal with $x_n = \xi n$, with $1 > \xi \geq c_1\epsilon_{\mathcal{N}}$. The second inequality in (4.17) fails, and Nagaev's lower bound cannot be applied. We choose $\delta > 0$ small enough so that $\xi(1+\delta)^2 < 1$, and we consider the event $\mathcal{A} = \{\forall z \in B(r_n), (1-\zeta_z) \geq \xi(1+\delta)^2\} \cap \{\tau_n > n\}$. Note that

$$\mathcal{A} \subset \left\{\sum_{z\in\mathbb{Z}^d} l_n(z)(1-\zeta_z(l_n(z)) \geq \xi(1+\delta)^2 n\right\}.$$

However, there might be some sites of $B(r_n)$ that the walk visits once, and if $\eta \in \{-1,1\}$, we will have on this sites that $\zeta_z(l_n(z)) = 0$. We will restrict to sites of $B(r_n)$ visited often. Note that, for $\alpha(\xi) > 0$,

$$\lim_{n\to\infty} Q\left(1-\zeta_z(n) \geq \xi(1+\delta)^2\right) = \lim_{n\to\infty} Q\left(\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n \eta_z(i)\right)^2 \leq 1-\xi(1+\delta)^2\right) = \alpha(\xi).$$

Thus, there is n_1 (depending on ξ and δ) such that for $n \geq n_1$

$$Q\left(\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_z(i)\right)^2 \leq 1-\xi(1+\delta)\right) \geq \frac{1}{2}\alpha(\xi).$$

Now, with n_1 fixed, we define a set

$$\mathcal{G}_n = \{z \in B(r_n):\ l_n(z) \geq n_1\}.$$

On the event $\{\tau_n > n\}$, we have for n large enough (using that $|B(r_n)| \ll n$)

$$l_n(\mathcal{G}_n^c) \leq |B(r_n)|n_1 \Longrightarrow l_n(\mathcal{G}_n) \geq n - |B(r_n)|n_1 \geq \frac{n}{1+\delta}.$$

Thus,

$$\mathcal{A} \subset \left\{\sum_{z\in\mathcal{G}_n} l_n(z)(1-\zeta_z(l_n(z)) \geq l_n(\mathcal{G}_n)\xi(1+\delta)^2 = \xi(1+\delta)n\right\}$$

Using (4.12) (with δ occurring in (4.12)), we have

$$\begin{aligned}2P\left(\sum_{z\in\mathbb{Z}^d} Y_z \geq \xi n\right) + \mathbb{P}_0\left(\mathcal{B}_n^c\right) \geq & (\frac{\alpha(\xi)}{2})^{|B(r_n)|} \times \mathbb{P}_0\left(\tau_n > n\right) \\ \geq & (\frac{\alpha(\xi)}{2})^{|B(r_n)|} \times \exp\left(-c_0\frac{n}{|B(r_n)|^{2/d}}\right).\end{aligned} \tag{4.22}$$

Since $1 > \xi > c_1\epsilon_{\mathcal{N}}$, the power of ξ appearing in (4.22) is irrelevant. We only need to check that the speed exponent is correct in (4.22)

4.3 The case $d \geq 4$ and $n^{\frac{d+2}{d+4}} \ll \xi_n \ll n$

Here $x_n = \xi_n$. Assume that we localize the walk a time T_n inside $B(r_n)$. We make use of Section 4.1 until the point where we assumed $T_n = n$ (that is a paragraph before (4.13)). If we were allowed to identify the two costs in (4.19), we would find here $T_n = x_n = \xi_n$, and $|B(r_n)| = \xi_n^{\zeta_d}$, with $\zeta_d = \frac{d}{d+2}$. Note that in dimension 4 or larger, with T_n of order ξ_n, we are not entitled to use Nagaev's lower bound. On the other hand, $|B(r_n)| = \xi_n^{\zeta_d}$, is the expected speed, so that constraining the local charges on $\mathcal{G}_n$ would yield the correct cost. We observe that we are entitled to use the CLT for $\zeta_z(l_n(z))$, for each sites in $\mathcal{G}_n$, since $l_n(z) \geq l_{T_n}(z) \geq \xi_n^{1-\zeta_d}$. With the notation Z for a standard gaussian variable, and n large enough, we have for $z \in \mathcal{G}_n$, and uniformely over $l_n(z)$

$$\alpha_0 := \frac{1}{2}P(Z^2 < \frac{1}{2}) \leq Q(\zeta_z(l_n(z)) < \frac{1}{2}).$$

With the choice $T_n = \frac{4}{\epsilon_0}\xi_n$ (note that $T_n \ll n$ for n large), recalling the definition of $\mathcal{G}_n$ in (4.5), and using that $l_n(z) \geq l_{T_n}(z)$

$$\left\{\forall z \in \mathcal{G}_n, \zeta_z(l_n(z)) < \frac{1}{2}\right\} \cap \left\{|\mathcal{G}_n| \geq \frac{\epsilon_0}{2}|B(r_n)|\right\} \subset \left\{\sum_{z\in\mathcal{G}_n} Y_z \geq \frac{1}{2}|\mathcal{G}_n|T_n = (1+\delta)\xi_n\right\}.$$

Thus, using (4.12)

$$2P\left(\sum_{z\in\mathbb{Z}^d} Y_z \ge \xi_n\right) + \mathbb{P}_0\left(\mathcal{B}_n^c\right) \ge \alpha_0^{|B(r_n)|} \times \mathbb{P}_0\left(\tau_n > T_n\right) \ge \exp\left(-c_d^- \xi_n^{\zeta_d}\right). \tag{4.23}$$

4.4 The case $d \ge 4$ and $x_n = \xi n$

We assume that $\xi < 1$, for $\delta' > 0$ so small that $(1+\delta')\xi < 1$, we choose $T_n = (1+\delta')\xi n$ and $|B(r_n)| = (\xi n)^{d/(d+2)}$. We force the local charges to realize $1 - \zeta_z(l_n(z)) \ge 1 - \frac{\delta'}{4}$ for δ' arbitrarily small. Note that for $\alpha_1 > 0$,

$$\lim_{n\to\infty} Q\left(\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n \eta_z(i)\right)^2 \le \frac{\delta'}{4}\right) = \alpha_1.$$

Thus, there is n_1 (depending on ξ and δ') such that for $n \ge n_1$

$$Q\left(\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n \eta_z(i)\right)^2 \le \frac{\delta'}{4}\right) \ge \frac{1}{2}\alpha_1. \tag{4.24}$$

Now, using n_1, we define a set

$$\mathcal{G}_n = \left\{z \in B(r_n) :\ l_n(z) \ge n_1\right\}.$$

On the event $\{\tau_n > (1+\delta')\xi n\}$, we have for n large enough (using that $|B(r_n)| \ll n$)

$$l_n(\mathcal{G}_n^c) \le |B(r_n)| n_1 \Longrightarrow l_n(\mathcal{G}_n) \ge (1+\delta')\xi n - |B(r_n)| n_1 \ge (1+\delta')(1-\frac{\delta'}{4}) n\xi.$$

We use (4.24) for $\zeta_z(l_n(z))$, with $z \in \mathcal{G}_n$. Thus, on $\{\tau_n \ge (1+\delta')\xi n\}$,

$$\{\forall z \in \mathcal{G}_n, \zeta_z(l_n(z)) < \delta'\} \subset \left\{\sum_{z\in\mathcal{G}_n} Y_z \ge (1-\frac{\delta'}{4}) l_n(\mathcal{G}_n) \ge (1+\delta')(1-\frac{\delta'}{4})^2 n\xi\right\}.$$

Now, we choose δ' so small that $(1+\delta')(1-\frac{\delta'}{4})^2 \ge 1+\delta$, for δ occurring in (4.12). Thus, using (4.12)

$$\begin{aligned} 2P\left(\sum_{z\in\mathbb{Z}^d} Y_z \ge \xi_n\right) + \mathbb{P}_0\left(\mathcal{B}_n^c\right) &\ge (\frac{\alpha_1}{2})^{|B(r_n)|} \times \mathbb{P}_0\left(\tau_n > T_n\right) \\ &\ge (\frac{\alpha_1}{2})^{|B(r_n)|} \times \exp\left(-c_0 \frac{(1+\delta')\xi n}{|B(r_n)|^{2/d}}\right) \\ &\ge \exp(-c_d^- (\xi n)^{\frac{d}{d+2}}). \end{aligned} \tag{4.25}$$

This yields the desired bound.

4.5 The case $d \geq 4$ and $n^{2/3} \ll \xi_n \ll n^{(d+2)/(d+4)}$.

The strategy in this region (region I of [4]) consists in letting the walk roam freely, while the *local charges* perform a moderate deviations. Note that our scenery ζ_z depends on the local times, and on sites visited only once by the walk, Y_z may vanish by (2.6), as in the model where $\eta \in \{-1, 1\}$. Thus, we only consider sites where $\{z : l_n(z) = 2\}$, since $\frac{1}{2}(\eta_1+\eta_2)^2-1$ is not degenerate. Also, a transient random walk has enough sites of this type. Indeed, Becker and König in [5] have shown that, in $d \geq 3$ with $\mathcal{D}_n(k) = \{z : l_n(z) = k\}$ for integer k, we have

$$\lim_{n\to\infty} \frac{E\left[|\mathcal{D}_n(k)|\right]}{n} = \gamma_0^2(1-\gamma_0)^{k-1}, \quad \text{where} \quad \gamma_0 = \mathbb{P}_0(S(k) \neq 0,\ \forall k > 0). \tag{4.26}$$

We choose a scenario based only on $\mathcal{D}_n(2)$. Note that for n large enough, the fact that $|\mathcal{D}_n(2)| \leq n$, and (4.26) imply that

$$\frac{1}{2}\gamma_0^2(1-\gamma_0) \leq \frac{E\left[|\mathcal{D}_n(2)|\right]}{n} \leq \mathbb{P}_0\left(\frac{|\mathcal{D}_n(2)|}{n} \geq \frac{1}{4}\gamma_0^2(1-\gamma_0)\right) + \frac{1}{4}\gamma_0^2(1-\gamma_0).$$

Thus,

$$\mathbb{P}_0\left(\frac{|\mathcal{D}_n(2)|}{n} \geq \gamma_1\right) \geq \gamma_1 \quad \text{with} \quad \gamma_1 = \frac{1}{4}\gamma_0^2(1-\gamma_0). \tag{4.27}$$

Now, we consider the following decomposition, for $\delta > 0$ small (recall that here $x_n = \sqrt{n}\,\xi_n$)

$$\left\{\sum_{z\in\mathbb{Z}^d} Y_z^{(n)} \geq \sqrt{n}\,\xi_n\right\} \supset \left\{\sum_{z\in\mathcal{D}_n(2)} Y_z^{(n)} \geq (1+\delta)\sqrt{n}\,\xi_n\right\} \cap \left\{\sum_{z\notin\mathcal{D}_n(2)} Y_z^{(n)} \geq -\delta\sqrt{n}\,\xi_n\right\}. \tag{4.28}$$

We treat the second event on the right hand side of (4.28) as in Section 4.1: we restrict to $\mathcal{B}_n$ (where $P(\mathcal{B}_n^c)$ is negligible by Corollary 2.9), and we use Markov's inequality.

Now, fixing a realization of the walk, $\{Y_z,\ z \in \mathcal{D}_n(2)\}$ are centered i.i.d with $E[Y_z^2] = 2(E_Q[\eta^4]+1)$, and on $\{|\mathcal{D}_n(2)| > \gamma_1 n\}$, then $\{\sum_{\mathcal{D}_n(2)} Y_z \geq (1+\delta)\sqrt{n}\,\xi_n\}$ is a moderate deviations. Thus, there is a constant $\underline{c}$, such that on the event $\{|\mathcal{D}_n(2)| > \gamma_1 n\}$, and for n large

$$\begin{aligned} Q\left(\sum_{\mathcal{D}_n(2)} Y_z \geq (1+\delta)\sqrt{n}\,\xi_n\right) &\geq \underline{c}\exp\left(-\frac{((1+\delta)\xi_n)^2 n}{2|\mathcal{D}_n(2)|(E_Q[\eta^4]+1)}\right) \\ &\geq \underline{c}\exp\left(-\frac{(1+\delta)^2\xi_n^2}{2\gamma_1(E_Q[\eta^4]+1)}\right). \end{aligned} \tag{4.29}$$

After integrating (4.29) the walk's law, we have

$$P\left(|\mathcal{D}_n(2)| > \gamma_1 n,\ \sum_{\mathcal{D}_n(2)} Y_z \geq (1+\delta)\sqrt{n}\,\xi_n\right) \geq \underline{c}\gamma_1\exp\left(-\frac{(1+\delta)^2}{2\gamma_1(E_Q[\eta^4]+1)}\xi_n^2\right). \tag{4.30}$$

5 Proof of Proposition 1.7

Large β First, $H_n \geq -n$ implies the upper bound in (1.18). The lower bound in (1.18) follows from the lower bound in (1.9) with $\xi_n = \xi n^{1/3}$, and the following inequalities: for $\xi < 1$

$$
\begin{aligned}
Z_n^-\left(\frac{\beta}{n^{2/5}}\right) =& E\left[\exp\left(-\beta\frac{H_n}{n^{2/5}}\right)\right] \geq P(H_n \leq -\xi n)e^{\beta\xi n^{3/5}} \\
\geq& \exp\left(n^{3/5}(\beta\xi - c_3^-\xi^{4/5})\right).
\end{aligned}
\tag{5.1}
$$

For any fixed $\xi < 1$, we choose β large enough so that the lower bound in (1.18) holds.

Now, define

$$
\mathcal{A}_n(a) = \left\{ |\{z \in \mathbb{Z}^d : \frac{n^{\frac{2}{5}}}{a} \leq l_n(z) \leq an^{\frac{2}{5}}\}| \geq \frac{n^{3/5}}{a^4} \right\}.
$$

Using the estimates of Section 3.1.2, we have for $\chi > 0$

$$
E\left[\exp(\left(-\beta\frac{H_n}{n^{2/5}}\right)\right] \leq e^{\beta n^{3/5}} P(\mathcal{A}_n^c(a)) \leq e^{n^{3/5}(\beta - \chi a^{2/3})}. \tag{5.2}
$$

Choosing a large enough so that $2\beta < \chi a^{2/3}$, and using the lower bound in (5.1), we obtain (1.19).

Small β**.** First, we decompose the partition function over the three regimes for $-H_n$: the moderate deviation, the large deviation and intermediate regimes. Thus,

$$
Z_n^-(\frac{\beta}{n^{2/5}}) = Z_I(\beta) + Z_{II}(\beta) + Z_{III}(\beta), \tag{5.3}
$$

with for ϵ small

$$
Z_I(\beta) = E\left[\exp\left(-\beta\frac{H_n}{n^{2/5}}\right) \mathbb{1}\left\{n^{\frac{1}{2}+\epsilon} < -H_n < n^{\frac{2}{3}+\epsilon}\right\}\right],
$$

$$
Z_{II}(\beta) = E\left[\exp\left(-\beta\frac{H_n}{n^{2/5}}\right) \mathbb{1}\left\{n^{\frac{2}{3}+\epsilon} < -H_n < n\right\}\right],
$$

and $Z_{III}(\beta)$ correponds to the remaining regimes.

We first deal with $Z_I(\beta)$ and rely on Chen's result (1.2). We note that from Chen's proof, his asymptotic result of (1.2) is actually uniform in the sequence ξ_n, in the sense that there is a sequence $\{\delta_n\}$ going to 0, such that for any $\xi_n \in [n^\epsilon, n^{1/6-\epsilon}]$, we have

$$
P(\frac{-H_n}{\sqrt{n}} > \xi_n) = \exp\left(-\frac{\xi_n^2}{2c_d}(1+\delta_n)\right). \tag{5.4}
$$

We have

$$
\begin{aligned}
Z_I(\beta) =& \exp(\beta n^{1/10+\epsilon}) + \beta\int_{n^{1/10+\epsilon}}^{n^{4/15-\epsilon}} e^{\beta u} P\left(\frac{-H_n}{n^{2/5}} > u\right) du \\
=& \exp(\beta n^{1/10+\epsilon}) + \beta n^{1/10}\int_{n^\epsilon}^{n^{1/6-\epsilon}} \exp\left(\beta n^{1/10}u - \frac{u^2}{2c_d}(1+\delta_n)\right) du
\end{aligned}
\tag{5.5}
$$

Now, the aymptotic behaviour is found as we maximize $\beta n^{1/10}u - \frac{u^2}{2c_d}$, which is $c_d\beta^2 n^{1/5}/2$. In other words, it is a simple computation that we omit, which yields for any $\beta > 0$,

$$\lim_{n\to\infty} \frac{1}{n^{1/5}} \log Z_I(\beta) = \frac{c_d\beta^2}{2}. \tag{5.6}$$

We deal now with Z_{II}, which corresponds to regime studied in Theorem 1.1. We will show that for β small, $Z_{II}(\beta) \le \exp(\epsilon n^{1/5})$, for ϵ small. Note that

$$Z_{II}(\beta) \le \sum_{k=0}^{\log_2(n^{1/3})} e^{2^{k+1}n^{4/15+\epsilon}\beta}\; P\left(2^k n^{4/15+\epsilon} \le \frac{-H_n}{n^{2/5}} < 2^{k+1}n^{4/15+\epsilon}\right) \tag{5.7}$$

In view of (5.7), it is enough to show that for $n^{3/5} \ge \xi_n \ge n^{4/15+\epsilon}$, we have

$$P(-H_n \ge \xi_n n^{2/5}) \le e^{-2\beta\xi_n}. \tag{5.8}$$

From (1.9), we have in this regime

$$P(-H_n \ge \xi_n n^{2/5}) \le \exp\left(-c_3^+ \left(\xi_n n^{2/5-2/3}\right)^{4/5} n^{1/3}\right), \tag{5.9}$$

and (5.8) requires that

$$c_3^+\xi_n^{4/5} n^{3/25} \ge 2\beta\xi_n \Longleftrightarrow \xi_n \le \left(\frac{c_3^+}{2\beta}\right)^5 n^{3/5}. \tag{5.10}$$

Since $\xi_n \le n^{3/5}$, (5.10) holds if $\beta < c_3^+/2$.

Finally, we deal with Z_{III}.

$$\begin{aligned} Z_{III} \le& \exp\left(\beta n^{1/2-2/5+\epsilon}\right) + \exp(\beta n^{2/3-2/5+\epsilon})P(-H_n \ge n^{2/3-\epsilon}) \\ \le& \exp\left(\beta n^{1/10+\epsilon}\right) + \exp\left(-\frac{n^{1/3-\epsilon}}{4c_d} + \beta n^{4/15+\epsilon}\right). \end{aligned} \tag{5.11}$$

Z_{III} is negligible when ϵ is such that $\frac{4}{15} + 3\epsilon \le \frac{1}{3}$.

We finally show (1.21). We choose $p > 1$ such that $p\beta < \beta_1$, and use Hölder's inequality

$$\begin{aligned} E\left[e^{-\beta\frac{H_n}{n^{2/5}}}\ 1\!\!1_{\{l_n(z)>bn^{1/5}\}\neq\emptyset}\right] \le& \left(E\left[e^{-p\beta\frac{H_n}{n^{2/5}}}\right]\right)^{1/p} \left(P(\exists z,\ l_n(z) > bn^{1/5})\right)^{1/q} \quad (q = \frac{p}{p-1}) \\ \le& e^{C\beta^2 n^{1/5}} \left(nP_0(l_n(0) > bn^{1/5})\right)^{1/q} \\ \le& n^{1/q}\exp\left((C\beta^2 - \frac{\chi_d b}{q})n^{1/5}\right). \end{aligned} \tag{5.12}$$

As we choose b large enough in (5.12), we obtain (1.21).

References

[1] Asselah, A., *Annealed Upper tails for the energy of a polymer.* Preprint.

[2] Asselah, A., *Shape transition under excess self-intersection for transient random walk*, To appear in Annales de l'Institut H.Poincaré.

[3] Asselah, A., *Large Deviations for the Self-Intersection Times for Simple Random Walk in dimension 3.* Probability Theory & Related Fields, 141(2008), no. 1-2, 19–45.

[4] Asselah, A., Castell F., *Random walk in random scenery and self-intersection local times in dimensions* $d \geq 5$. Probability Theory & Related Fields, 138 (2007), no. 1-2, 1–32.

[5] Becker M.; König W. , *Moments and distribution of the local times of a transient random walk on* $\mathbb{Z}^d$. Preprint 2007.

[6] Biskup M.; König W. *Long-time tails in the parabolic Anderson model with bounded potential.* Ann. Probab. 29 (2001), no. 2, 636–682.

[7] Buffet, E; Pulé J.V., *A model of continuous polymers with random charges* J. Math. Phys. 38, 5143 (1997)

[8] Chen, Xia *Limit laws for the energy of a charged polymer.* To appear in Annales de l'I.H.Poincaré, 2008.

[9] Chen Xia, *Random walk intersections: Large deviations and some related topics* book in preparation 2008.

[10] Chen Xia, Khoshnevisan Davar, *From charged polymers to random walk in random scenery*, preprint 2008.

[11] Derrida B., Griffith R.B., Higgs P.G. *A model of directed walks with random interactions.* Europhysics Letters, 18 (1992), 361-366.

[12] Derrida B.; Higgs P. G. *Low-temperature properties of directed walks with random self-interactions.* J. Phys. A 27 (1994), no. 16, 5485–5493.

[13] Gantert N.; van der Hofstad R.; König W. *Deviations of a random walk in a random scenery with stretched exponential tails.* Stochastic Process. Appl. 116 (2006), no. 3, 480–492.

[14] Garel T.; Orland H. *mean field model for protein folding* Europhysics Letters, 6(1988), 307.

[15] Kantor Y., Kardar M. *Polymers with self-interactions.* Europhysics Letters, 18 (1992), 14,(1991),421-426.

[16] Lawler G., Bramson M., Griffeath D. *Internal diffusion limited aggregation* The annals of Probability, vol 20, (1992), no.4, 2117–2140.

[17] Nagaev, A. *Integral limit theorems for large deviations when Cramer's condition is not fulfilled* (Russian)I,II Teor. Verojatnost. i Primenen. 14 (1969) 51–64, 203–216.

[18] Nagaev, S. *Large deviations of sums of independent random variables* Annals of Probability,7(1979),no.5, 745–789.

[19] Nagaev, S. *Lower bounds for the probabilities of large deviations of sums of independent random variables.* (Russian) Teor. Veroyatnost. i Primenen. 46 (2001), no. 1, 50–73; translation in Theory Probab. Appl. 46 (2002), no. 1, 79–102